

Transparency to Law-Drafting: Lawradar as Every Finns' Own Watchdog

Presented at the ECPR General Conference 2024, University College Dublin, 12 – 15 August 2024

István Rytkönen, Marjaana Sjölund
University of Turku

Abstract

Lawradar (accessible at <https://lakitutka.fi/>) is a ground-breaking policy research tool for law-drafting documents in Finland. It pools together all official documents produced during the legislative process of individual bills, ranging from the preliminary preparation phase to enactment. At the core of its novel interface, '*Lawradar for All*', is the idea that the user might not have any prior knowledge about the policy processes behind legislation, yet they will be able to study all topics of their interest using Lawradar. Lawradar makes legislative documents truly accessible for both researchers and citizens alike, aiming to democratise information that impacts everyone.

During the latest development phase of Lawradar, we have focused on users without an academic background in law or politics. Hence, the simplicity of language and the unambiguity of the legislative process are the focal points in the development. The previous word-based search is augmented with a semantic search that relieves the users from using overly formal or technical language. A new massive open online course is also being developed by our team and linked tightly with Lawradar. The course will offer everyone an insight into the legislative process and the opportunities of participation in state-level legislative processes – and how to stay informed about proposed law changes affecting their lives.

Furthermore, the Lawradar research team continues to improve the usability of legislative documents for all. We envision creating a legislative reform focused chatbot to answer any questions related to past and present legislative processes, and to facilitate the targeted and detailed consultation of all relevant stakeholders and individuals. This chatbot would be able to summarize the effects the proposed law would have to the particular user chatting with the bot – and even facilitate the submission of their feedback to the legislator. The advent of LLM-based solutions is providing us the necessary technology to aid both the citizens in following and participating in the process, and the government officials in consulting the stakeholders. A legislative reform focused chatbot would be a hopeful step towards a more transparent and inclusive legislative process with the help of modern technology.

Keywords: Democracy, Policy Analysis, Political Participation, Decision Making, Lobbying, Technology, Big Data, Policy-Making

Introduction

In Finland, formal participation in the law-drafting process is an opportunity commonly provided to stakeholders before a bill is passed to the parliament from the ministry in charge. However, the formal and technical nature of the process, as well as the demandingly vast amount of information to be analysed before participation, pose barriers to the idea of equal participatory opportunities for all stakeholders.

In the context of public policy-making, the term 'stakeholder' refers to the different parties that are affected by the decisions (Sara 2018). This includes public servants as well as public and private organisations and private persons. Participation of these stakeholders in law-making is most often considered to be an asset for both the quality and legitimacy of laws in democratic societies (Crow, Albright & Koebele 2019). What is more, the knowledgeability of the individuals and experts participating in the

process is found to make a crucial difference for the quality of legislation (Hong 2015). In addition, more transparency in the process and better participation opportunities for stakeholders balance the unequal power relations in law-making, and force the focus to be more on the public interest than on specific powerful private interests (Keinänen & Paasonen 2015, p. 6).

The Finnish law-making process is divided into two successive parts as seen in Figure 1: First comes the drafting of the laws and their preparatory materials in the ministries, where the ministry officials are in charge of the process. Second part is the actual decision-making, which is done by the members of the Finnish parliament, and is based on the information gathered in the previous phases.



Figure 1 Overview of the Finnish law-making process

With the exception of expert hearings in the Finnish parliament, the early law-making phases in ministries provide the last opportunity for stakeholders and individuals to have an actual impact on the contents and form of a piece of legislation. Although the actual legislative power has been vested to the parliament in the Finnish constitution, the contents of the laws very seldom change in the decision-making phase (Tala 2001, pp. 25, 105; Pakarinen 2011, p. 61), rendering participation in these later phases ineffective. Thus, the consultation phase constitutes a crucial element in the democratic apparatus of Finland.

Organisatory stakeholders in Finland have identified various shortcomings in the participative nature of the law-making process. Only 28 % of stakeholders answer that they deem officials to be successful in taking marginalized, “silent” actors into consideration when they draft bills that concern these actors. Furthermore, only 36 % of the respondents agreed with the statement that the methods for carrying out consultations are versatile. (Uusikylä et al. 2023.) The need for more inclusive, tailored consultations has been recognised with particular attention to silent groups (SILE 2023).

Natural language processing (NLP) solutions have been identified at the University of Turku as ideal candidates for bridging gaps in democracy and transparency in Finnish law-making. One proposed solution utilising NLP is a legislative reform focused chatbot built to answer detailed questions and explain complex legislative phenomena in laymen terms. This kind of modern chatbot has the potential to drastically reform the consultation of both the organisatory stakeholders and individuals during law-making.

Lawradar as a method of educating ‘all’ about law-making in Finland

Lawradar is a Finnish open-access research infrastructure that pools together the official documents produced during the legislative processes from the preliminary preparation phase to enactment. It is a pioneer in its field, both nationally and internationally, as it serves as the crucial link that connects legislative documents together and makes the information easily accessible online. Lawradar combines documents from multiple sources, including data from governmental open-access APIs and digitalised documents, but also paper documents archived in the National Archives of Finland that have never been digitalised before. The digital tools of the infrastructure are designed to expedite and enhance the research of law-making, revealing the legislative footprint and promoting unprecedentedly broad research on the interrelations between the legislator, stakeholders, courts, and scholars.

Having been launched in 2020, Lawradar has since evolved from a research infrastructure into an intuitive and instructive tool for transparency and accountability, maintaining nevertheless the vast array of precise and reliable search and analysis functions.

Although Lawradar has always strived to be an easy-to-use interface to access all legislative documents, it was originally targeted towards a segment of people somewhat familiar with the legislative process, e.g. political journalists, law students, lawyers, and political, socio-legal and legal scholars. During the beginning of the SILE project (“Silent Agents Affected by Legislation”), it became clear that not everyone is equally able to follow the law-making processes. To even the field, ideas of various usability improvements emerged, with which Lawradar could serve as an education platform for ‘all’. Thus, for the latest development project the mission has been to reach ever more audiences that do not hold prior understanding of the legislative process, including those with potential grudges and frustration towards the complexity of law-drafting and its limited elements of direct democracy. In addition to serving researchers and other professions, Lawradar also improves the access to information for the general public – including the silent actors – thus supporting a fundamental right and a cornerstone of democracy.

The new user interface that will supplement the previous researcher-oriented interface, is illustrated in Figure 2. Some of the new usability features that help everyone to use Lawradar as a source of information are emphasised. First, we predefined some criteria to searches we see as interesting, topical, or otherwise important, and added direct links to them on the frontpage. Second, the new user interface has a single search field to cover all document types, and it utilises the semantic search described later in section 3.2. Third, an automatically updating word cloud, implemented using accessible theme bubbles, represent the most relevant themes for the past year and hopefully inspire the visitor to browse new topics. Fourth, we added links to videos and instructive text snippets about law-making to help users understand the contents. The changes from the existing researcher-oriented interface are notable.

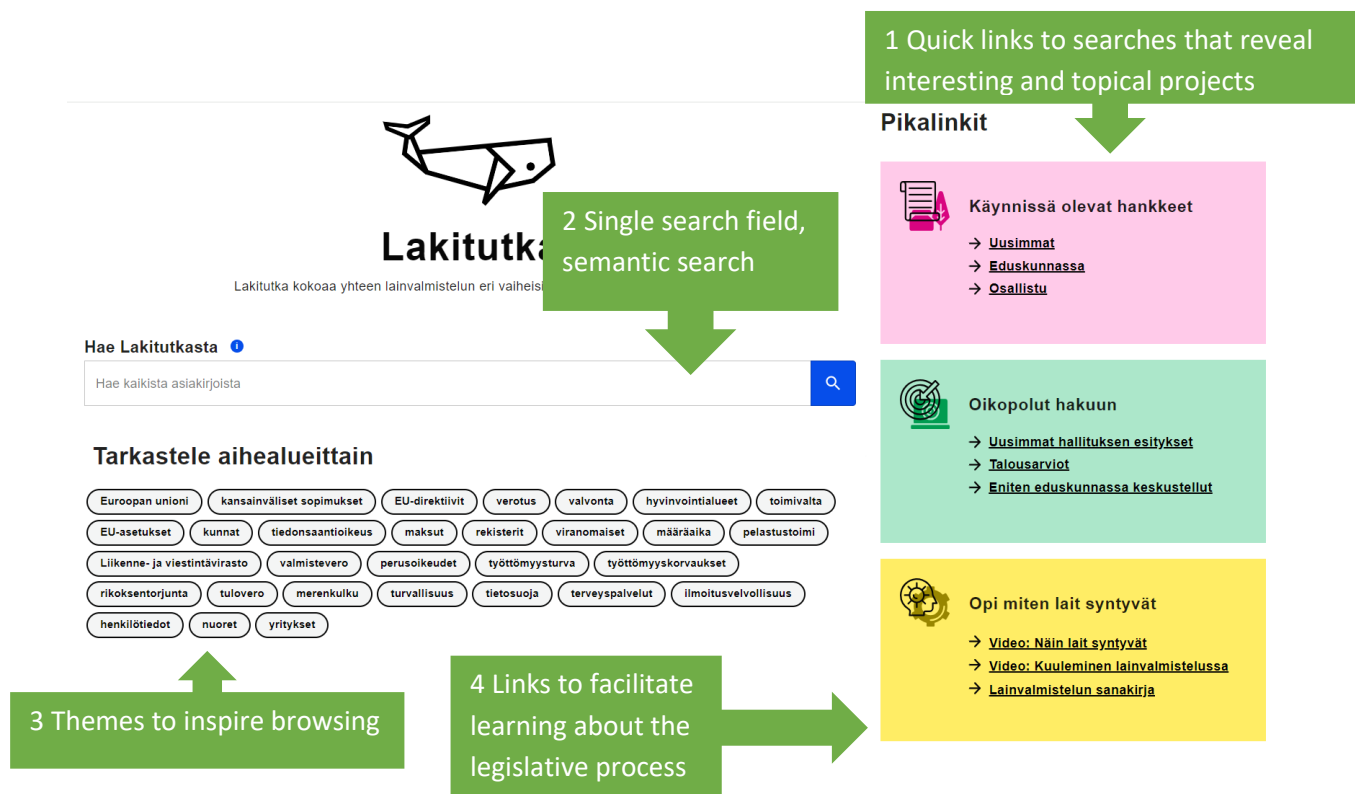


Figure 2 Lawradar for All

In addition to the improvements in the user interface of Lawradar, the idea of 'Lawradar for All' has been implemented in other ways, too. Firstly, a massive open online course (MOOC), affectionally called 'The Law Drafting School', is in final stages of development by our team. The MOOC will offer everyone an insight into the legislative process and educate them about their own opportunities to participate in state-level legislative processes – and how to stay informed about proposed law changes affecting their lives. It can be studied as a stand-alone course, but it will also be linked tightly with Lawradar to serve as a user manual. Secondly, the access to Lawradar has been granted to prisoners inside Finnish prisons (Rikosseuraamuslaitos n.d.), who have a tightly restricted access to the internet. Thirdly, in the context of Lawradar, 'All' includes also the Swedish-speaking minority living in Finland, as well as any users in Nordic countries. That is why the interface and the data – previously available only in Finnish – will be from now on available also in Swedish that is the other official language of Finland.

There is, however, a limit to how much transparency and democratic progress may be achieved using these kinds of user interface improvements to the existing Lawradar web service, and by merely improving the access to the documents. Hence, completely new solutions are also needed, and we have committed to expanding our focus from the tools to the content itself.

NLP in the context of law-drafting documents

Characteristics of textual data of legislative reforms

The legislative reforms are an important object of research especially in the sociological and socio-legal fields, but provide a lot of possibilities for political scientist, linguistics, and more. Although most of the documents are not legally binding in nature, lawyers use the *travaux* – preparatory documents – as evidence of the legislator's intent when making interpretations of the law (Alvesalo-Kuusi & Kumpula 2021, p. 45-46). For others, the *travaux* give insight into the power struggles behind legislative decisions: who has been heard about the content and form of the law, and whose voices have been ignored (see e.g. Nieminen & Sarasoja 2023)?

The data in Lawradar's data repository covers years between 1976 and 2024, and is comprehensive from 2015 onwards. The focus is on the preparatory documents, not on law text. There is a clear difference between the high formality and rule-baseness of the law texts, and the natural language utilised in everyday texts like news. Legislative documents situate themselves in the middle: some of them are quite formal and adhere to specific standards of writing, but a lot of the documents are completely freeform in nature.

The preparatory data include documents drafted in the ministries in the process of law-drafting, and documents written in the parliament during the decision-making part of the law-making process, with their associated metadata. The documents are of varying types, e.g. memos, reports, draft bills, bill proposals, statements, and of varying characteristics, which complicates their combined analysis. Firstly, the style of the text varies depending on the author. A large proportion of the documents has been drafted by government officials. For example, the legislative counsellors employed in the ministries write the drafts and bill proposals, and the committee counsellors are responsible for writing the committee reports. However, a large number of documents have been written by different stakeholders with different levels of familiarity with the legislative style of writing. These documents, typically statements that represent the stance of the stakeholder to a specific legislative reform, can vary significantly in their format, formality, etc. Secondly, although the language of the data is predominantly Finnish, a part of the data exists either only in Swedish, or both in Finnish and Swedish. According to the Language Act (423/2003), the law

proposals are typically written in Finnish, but in particular cases they are written in, or translated into, Swedish. The stakeholders have a right to submit their comments in either language. Thus, a multi-lingual approach is needed to cover all aspects of the dual-language legislative projects.

The amount of data grows quickly. Firstly, the data cumulates constantly as new legislative reforms are being discussed and proposed. Secondly, older documents are being digitalised as well as new related documents and metadata are being added to the repository. The aforementioned reasons cause the amount of data in the repository to grow to the extent that makes it unfeasible to comb through the data manually. According to the data in Lawradar data repository, in Finland the average number of legislative projects that result in the government making a law proposal is 238 per year (for 2000-2023), not counting the numerous projects that are abandoned during law-drafting, nor the proposals the individual members or the citizens make. Each project has an average of 35 documents, including draft versions, bill proposals, stakeholder comments, committee statements, etc, while a typical document is little over 14 000 characters in length (including all headers, footers, signatures, etc.). In total, the data repository currently contains textual data from machine-readable documents in extent of more than 3,2 billion characters. Therefore, we have begun to utilise NLP solutions to help with analysing and processing the data.

Semantically searching law-drafting documents and parliamentary debates

Semantic search as a term refers to techniques used for search functions which, in contrast with string matching, are designed to find relevant text by analysing its context and meaning besides the exact match of text strings (Kasenchak 2019). Such approaches have the potential to provide more relevant search results with a lesser number of user actions.

Lawradar has traditionally employed string-matching search functions, as the use of those provides certain reliability when the user is searching for exact terms and prefers not to include any results that relate to the search terms without exact match. However, in the last year, a parallel option has been developed for Lawradar: those not entirely certain about the exact search terms may make use of the semantic search function. The semantic search is based on a Finnish SentenceBERT model developed at the University of Turku (Kanerva et al. 2021). This model is trained on Finnish paraphrase corpus, making it an ideal candidate to detect quasi-paraphrases, or phrases that have approximate equivalence (Bhagat & Hovy 2013). In practice, this allows the user e.g. to use colloquial terms as they aim to find a set of documents related to any legislative process, and it may also save time and effort from more advanced users, such as researchers, in gathering all documents and data relevant to the phenomenon they are examining.

Table 1 Example use cases for semantic search

<i>Use cases</i>	String matching	Semantic search
<i>A. User with little prior knowledge searching information about legislation concerning fuel tax cuts</i>	fuel AND (tax OR taxation) diesel AND (tax OR taxation) ...	<i>cheaper gas</i>
<i>B. Researcher aiming to identify legislative changes that have affected energy prices in Finland</i>	energ* OR fuel OR electri* OR nuclear OR ...	<i>effects on energy prices</i>

As demonstrated in Table 1, a major use case A is about a layman user trying to find legislative changes based on their own everyday experience. They might be interested to see when and why the gasoline was

made cheaper (or more expensive, on the other hand), and are not concerned if the price was changed due to taxation changes or other changes in the law. For the string-based search, as the language of the bills is quite formal, the search for “cheaper gas” would probably only return a couple of search hits from the parliamentary speeches, and none from the legislative documents themselves. By utilising the semantic search, the search engine will be able to recognize “gas” as belonging semantically to the same group as “fuel”, “diesel”, and “cheaper” having a semantically similar meaning to “lowering taxes”.

The other major use case B is about the researcher who is doing preliminary data gathering, aiming to identify all of the legislative changes they want to include into their research. They could utilise the semantic search to broaden the “net” without needing to specify each search word separately. As the semantic search supplied by Lawradar does not require the phrases to have identical meaning, only a high level of similarity, the results are not directly suitable to act as the research data. That is why, for the generation of the research dataset, the researcher would then continue on selecting only the relevant search results as data.

While implementing the semantic search, we stumbled upon a challenge regarding what part of the documents the search should target. Our data units, the documents, consist of dozens or even hundreds of pages of text, while the search string the user gives is typically two or three words long. Locating the part of the documents that contain the essence of the document, and converting that to a vector representation, was integral part of fine-tuning the performance of the semantic search.

The uses of a chatbot to facilitate stakeholder participation

The Finnish law-making process usually allows anyone to participate for example by sending their comments through a service called *lausuntopalvelu.fi*. In reality, however, this participation opportunity is not accessible to a large proportion of the individuals and smaller associations, since properly commenting a draft requires understanding the content of the proposed reform. The reform might contain hundreds of pages of text, so it is very unlikely that the individuals whose lives the reform will affect will read even a small portion of it. They must instead rely on secondary sources, e.g. media outlets, social media influencers or politicians, which significantly hinders the potential of participatory law-making. Some individuals might be represented by interest groups and other stakeholders to a certain extent, but this cannot be considered a comprehensive and direct means of participation.

From the ministry officials’ perspective, a thorough consultation with traditional methods is a tedious task. Stakeholders most commonly submit opinion documents to the ministry, or answer a brief set of questions through an online consultation portal. Certain officials are known to make more direct contact with stakeholders, e.g. through phone calls, e-mails and surveys, as well as making use of more interactive consultation platforms, but these cases are far from being a general rule. The obvious limitations to the high-quality law-drafting, which includes the use of versatile hearing methods, are the shortage of time and personnel (Uusikylä et al. 2023, pp. 50-52). Both are needed for reaching out to stakeholders and individuals in an appropriate manner, and consequently for summarising and analysing the results of consultation.

A possible solution offered by the Lawradar team is a legislative reform focused chatbot. This bot would serve a dual purpose of summarising the essences of the reforms and enabling a direct line of feedback to the ministry officials. The legislative reform focused chatbot would not replace the ministry officials, but would offer completely new opportunities to keeping stakeholders informed and consulting them. The

chatbot could serve as an instructive assistant, explaining any contents of a draft bill, and, if opted so by the user, it could also ask tailored questions from the user to collect their views on the reform in question, providing the relevant ministry with novel and precise consultation data. The use of such assistant would entail a ground-breaking shift from a tedious, formal hearing process towards a truly open forum where all individuals have the fair chance of understanding the effects of legislative projects and participating at an extent suitable for them personally.

As the chatbot is still at the planning stage, nothing definitive can be said about the technologies to be employed. In the design phase, a multitude of necessary features have been identified: The bot needs access to both topical and historical legislative data, with the data accumulating all the time. The bot has to be able to summarise data from any legislative reform. The bot needs to understand written questions and reply to them correctly, and finally, be able to query the user for feedback.

The advent of Large Language Model (LLM) based solutions is providing us the necessary technology to perform all these in a one solution. Even if the LLMs are performing well in a multitude of tasks, pitfalls have also been identified. One of these is the fact that LLMs are known to hallucinate and create information out of thin air (Ji et al. 2023), which is unacceptable in this kind of knowledge-intensive task. The language model cannot be trained every time new information is added; hence this domain-specific task requires a model that can utilise up-to-date information instead of relying on trained data. A Retrieval Augmented Generation model is a possible solution to these problems. Combining the parametric memory standard with language models with a retrieval-based knowledge base have resulted in more factual language while simultaneously allowing the data to be updated without training the model from scratch. (Lewis et al. 2020.) In our case, the knowledge base, or the non-parametric memory, would be the text information stored in our database about the legislative reforms. As we have not proceeded yet to the implementation phase, the suitability and performance of this kind of model is still uncertain.

Conclusions

In the coming years, Lawradar will focus on taking concrete steps towards enhanced inclusion and participation in the law-drafting process in Finland – and monitoring the true effects of these. This will include particular attention to e.g. the unambiguity of language, explanatory information bubbles guiding the use of the services, more visual and clickable features, and the flexibility of the semantic search.

When the legislative reform focused chatbot moves to the actual implementation stage – either by us, or by Finnish governmental authorities – it is a concrete step towards a more transparent and inclusive legislative process with the help of modern technology.

Funding

Lawradar for All -project has been funded by the Strategic Research Council (decisions 335442/335654 and 358263/358264) with additional funding from the European Union's NextGenerationEU instrument (Academy of Finland grant numbers 352827, 353569 and 352828).

References

- Alvesalo-Kuusi, A. & Kumpula, A. (2021). "Oikeussosiologia, oikeustiede ja lainoppi". In Nieminen, K. & Lähteenmäki, N., *Empiirinen oikeustutkimus*. Gaudeamus.
- Bhagat, R. & Hovy, E. (2013). "Squibs: What Is a Paraphrase?". *Computational Linguistics* 39(3): 463–472.
- Crow, D. A., Elizabeth A. A., and Elizabeth A. K. (2019). "Stakeholder Participation and Strategy in Rulemaking: A Comparative Analysis." *State politics & policy quarterly* 19(2): 208–35.
- Hong, S. (2015). "Citizen Participation in Budgeting: A Trade-Off between Knowledge and Inclusiveness?" *Public administration review* 75(4): 572–82.
- Kanerva, J, Ginter, F., Chang, L., Rastas, I., Skantsi, V., Kilpeläinen, J., Kupari, H. et al. (2021). "Finnish Paraphrase Corpus." In *Proceedings of the 23rd Nordic Conference on Computational Linguistics*, Linköping University Electronic Press, Sweden, 288–98. <https://aclanthology.org/2021.nodalida-main.29>.
- Kasenchak, R. T. (2019). "What Is Semantic Search? And Why Is It Important?" ed. Bonnie Lawlor. *Information Services & Use* 39(3): 205–13. doi:10.3233/ISU-190045.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S. & Kiela, D. (2020). "Retrieval-augmented generation for knowledge-intensive NLP tasks". *Advances in Neural Information Processing Systems*, 33, 9459–9474.
- Nieminen K. & Sarasoja L. (2023) "Epistemic othering: the interplay of knowledges in legislative drafting". *Journal of Law and Society* 50: 322–343. <https://doi.org/10.1111/jols.12443>
- Pakarinen, A. (2011). "Lainvalmistelu vuorovaikutuksena: analyysi keskeisten etujärjestöjen näkemyksistä lainvalmisteluun osallistumisesta". Oikeuspoliittisen tutkimuslaitoksen tutkimuksia 253. Helsinki: Oikeuspoliittinen tutkimuslaitos.
- Rikosseuraamuslaitos (n.d.). "Yhteiskunta." https://www.rikosseuraamus.fi/fi/index/portti/yhteiskunta_0.html (Accessed June 4, 2024).
- Sara, I. (2019). "Whose voice? Understanding stakeholder involvement in law drafting affecting Sámi reindeer herding". JYU dissertations 44. University of Jyväskylä.
- SILE (2023). "Hiljaiset ryhmät huomioitava paremmin lainvalmistelussa". <https://ratkaisujatieteesta.fi/osallistuva-kansalaisuus-ja-muuttuva-hallinta/hiljaiset-ryhmat-huomioitava-paremmiin-lainvalmistelussa/> (Accessed June 25, 2024).
- Tala, J. (2001). "Lakien vaikutukset: lakiuudistusten tavoitteet ja niiden toteutuminen lainsäädäntöteoreettisessa tarkastelussa". Oikeuspoliittisen tutkimuslaitoksen julkaisuja, 177:2001. Helsinki: Oikeuspoliittinen tutkimuslaitos.
- Uusikylä, P., Keinänen, A., Vartiainen, N., Ervasti, K., Salminen, V., Kettinen, J., Lintinen, U., Köppä, L., and Lindström, E. (2023). "Kohti laadukasta lainvalmistelua: Lainvalmisteluprosessin laatuindikaattorit". Valtioneuvoston kanslia.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). "Survey of hallucination in natural language generation". *ACM Computing Surveys*, 55(12), 1-38.